(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 061 487 A1

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
20.12.2000 Bulletin 2000/51

(51) Int Cl.7: G08B 13/194

(21) Application number: 99830376.2

(22) Date of filing: 17.06.1999

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(71) Applicant: ISTITUTO TRENTINO DI CULTURA
I-38100 Trento (IT)

(72) Inventors:
- Aste, Marco
38068 Rovereto (Trento) (IT)
- Boninsegna, Massimo
38037 Predazzo (Trento) (IT)

(74) Representative: Bosotti, Luciano et al
c/o JACOBACCI & PERANI S.p.A.
Corso Regio Parco, 27
10152 Torino (IT)

# (54) A method and device for automatically controlling a region in space

(57) The monitored region (S) is monitored by image signal generating means such as video cameras (2) in order to obtain a succession of images of the bodies (A, B) present in the monitored region, each image corresponding to a defined instant. The images are processed in such a way as to obtain, for each instant considered, a volumetric map of each body present in the region (S). This map, which identifies characteristics of shape, position, volume and dimensions of the body to which it refers, is processed in order to extract from it at least one parameter selected from the following group: descriptors of shape and volume, such as the volumetric map itself, the co-ordinates of position and the dimensions of each body to which the volumetric map refers. The parameter or a succession of values of the parameter obtained in this way is then compared with at least one model of these characteristics stored in a processing unit (1). Depending on the outcome of this comparison operation, a procedure of surveillance and/or reporting may be selectively activated. The solution is applicable, for example, to the automatic monitoring of museum environments, e.g. to ensure that visitors do not come too close to an exhibited work, or to the monitoring of industrial environments, e.g. to ensure that an operator does not come too close to a dangerous machine or process.

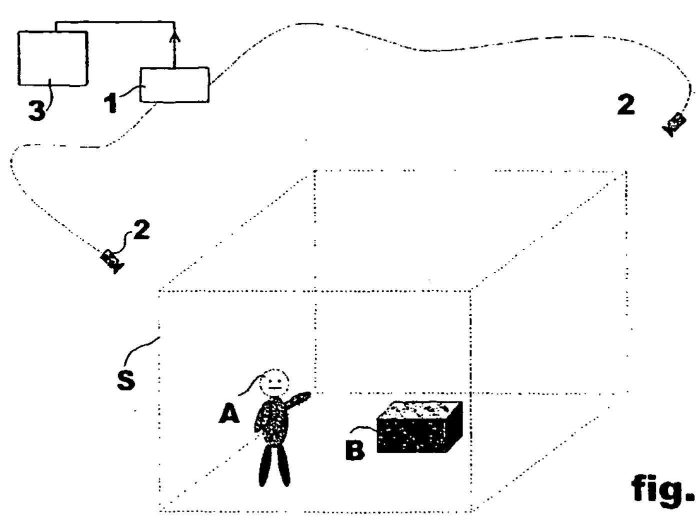



fig.1

EP 1 061 487 A1

## Description

[0001] The present invention relates in a general way to the automatic monitoring of a region of space, particularly as regards the detection and location of bodies present within this region.

[0002] The ability to detect and locate bodies present within a region of space is useful when it is necessary to monitor the presence of people or objects in particular regions or subregions.

[0003] For example, for reasons of security, the presence of people in defined areas may be judged to endanger the people themselves, or the objects present within that area. To take a particular case, in an industrial environment a volume of space in the vicinity of a machine that may produce chips or liquids dangerous to man; or the volume represented by the radius of action of a mechanical arm; or regions of space close to equipment operating at high voltages; or more generally any region in which machines dangerous to man are operating, may be considered dangerous. Hence the desirability of being able to automatically monitor and report the presence of people in a region of space regarded as dangerous for human activity.

[0004] Another example of an application is the protection of objects of artistic value in museum environments or in any situation in which the presence of people within a certain region can be a source of danger for the objects present in that area.

[0005] At the present time, regions of space are monitored automatically by devices such as optical barriers of transmission and/or reflection type (typically based on infrared technology), physical barriers, pressure-sensitive mats, movement detectors based on microwaves, passive infrared or ultrasound, radar systems, and devices that use laser beams to detect the presence and position of objects.

[0006] In many cases it is difficult to use these techniques for reasons of practicality, layout and reliability or environmental compatibility. Known systems, in fact, have intrinsic limitations which make their use difficult. For example, many known systems are sensitive to noise, dust and dirt and are therefore unsuitable for use in industrial working environments. Other limitations have to do with the difficulty of discriminating the size of the detected object and/or the inability to analyse the behaviour and motion of bodies in the vicinity of and within the monitored region of space. Again, some systems are unsuitable as being too invasive in environments which should be respected such as museums, or more generally places of great historical and artistic worth. Moreover, many of the known systems can easily be deceived, while others, such as physical barriers, may be unacceptable for reasons of safety and/or practicality; while others, such as those based on the emission of electromagnetic radiation, may not be tolerated because of their interference with other equipment, the difficulty of setting them up and in some cases the danger which they may present to biological organisms.

[0007] Many of the problems cited above can be overcome by monitoring the observed region with image signal generating means represented - in commonly used surveillance systems - by video cameras, sometimes of the type often known as "slow video". These systems however have the drawback that they require the constant presence of a human operator if they are to be of any real benefit.

[0008] The object of this invention is therefore to provide a solution for the automatic detection of bodies within a defined region that is simple, reliable, easily set up and capable of discriminating between objects for shape and volume while also considering the movement and relative path of the objects within the monitored area.

[0009] According to this invention, this object is achieved with a method having the characteristics claimed specifically in the following claims. The invention also relates to apparatus for carrying out this method.

[0010] In summary, the solution according to the invention is capable of automatically detecting, locating and reporting the presence of bodies within a monitored region of space. The invention is also capable of discriminating with a high degree of reliability between objects of different shapes and sizes and is therefore able to select, for monitoring purposes, only one particular type of body; e.g. only people.

[0011] In particular, the solution according to the invention is capable of detecting with a high degree of reliability the simultaneous presence of bodies, not necessarily of the same shape, in the monitored region, and selecting for monitoring purposes only those bodies that present certain distinctive features, for example only people or objects above a certain height. In addition, the quantity of information obtainable with the solution according to the invention is very much greater than could be obtained by means of conventional techniques of automatic detection, and makes possible more reliable and robust location of bodies within the monitored area, overcoming the limitations from which known systems usually suffer and so enabling it to be adapted more successfully to different working conditions and hence giving it greater generality of use.

[0012] In specific terms, the solution according to the invention is able to carry out a volumetric analysis and so extract the characteristics of form, volume and dimensions which distinguish an object or person which it is wished to pick out from other artefacts or objects that should be inside the monitored region. For example, in order to recognize the presence of a person, it is possible to use the a priori knowledge that a person possesses a certain shape and therefore produces a characteristic occupied volume.

[0013] With the invention it is therefore possible to discriminate between objects and people regardless of how they are moving and to use the information obtain-

able by the method in order to define different degrees of danger and/or alarm as a consequence of the presence of bodies within defined subregions of space.

[0014] The invention will now be described. purely by way of non-restrictive example, with reference to the attached drawings, in which:

- Figure 1 shows diagrammatically the characteristics of the system according to the invention used for the monitoring and automatic surveillance of a defined region of space,
- Figure 2, comprising four parts respectively labelled a1-a2 and b1-b2, illustrates the generation of image signals within the context of the solution illustrated in Figure 1,
- Figures 3 and 4 illustrate, in ways basically identical to those of Figures 1 and 2, another possible embodiment of the solution according to the invention; in particular, Figure 4 is composed of eight parts respectively labelled a1-a2, b1-b2, c1-c2 and d1-d2,
- Figure 5 illustrates the methods adopted for calculating a so-called map of volumetric occupation,
- Figure 6 illustrates schematically one of these maps capable of being obtained within the context of the invention, and
- Figure 7 is a flow diagram relating to the generation and use of such a map.

[0015] In particular the expression "volumetric map" as used here means any representation of occupied volumes due to the presence of a body, in other words a representation of a three-dimensional map in which the regions of volumetric occupation introduced by the presence of bodies are indicated. Such a map is obtained after image analysis procedures have been carried out using automatic methods known per se or according to the embodiments of the invention described below. For a summary of some of these methods the following may usefully be referred to: Marr D., "Vision", Freeman, 1982; Ballard D.H. and Brown C.M. "Computer Vision", Prentice Hall, 1982; Martin W.N. and Aggarwaal J.K., "Volumetric description of objects from multiple views", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 5, pp.150-158, 1983. From this map, by means of the volumetric analysis carried out using automatic methods known per se it is possible to derive the characteristics of shape, volume, dimensions and position of the bodies present in a defined region of space in such a way that they can easily be compared with similar representations obtained from the volumetric maps of other bodies.

[0016] In both Figure 1 and Figure 3, the reference S indicates a region of space in which it is wished to detect the presence of people A or objects B.

[0017] The region S may be bounded by physical walls , as for example in the case of a room or cage, or may consist simply of a portion of space bounded by an imaginary closed surface that separates a generic space in two regions, or it may be bounded partly by physical barriers, for example the floor, and partly by an imaginary surface. The monitored region has however the feature of a volume, may be of any shape and can be defined simply and flexibly according to need.

[0018] In the currently preferred embodiment of the invention, the volumes occupied by the bodies (such as bodies A, B visible in Figures 1 and 3) that are present in the monitored region S are found by using two or more video cameras (acting as image signal generating means) installed in such a way that the region S is in the visual field of at least two video cameras 2, as shown for example in Figures 1 and 3 (the latter figure referring to a solution in which four video cameras 2 are used). It is advisable for the video cameras 2 to be so positioned as to avoid occlusions due to the movement of objects on the same plane; for example, for bodies of different heights standing on the floor and moving about, it is preferable to have views from above.

[0019] The signals (of analogue type or already directly converted into digital form) output by the video cameras 2 are sent to a processing unit 1 which may be a specialized processor or, in the currently preferred embodiment of the invention, a computer such as a programmed personal computer (known per se) in order to extract from the images the shapes of the bodies A and B present within the region S to be monitored. The object here is to check for the possible presence of bodies not inherently belonging to the monitored region.

[0020] In particular, Figures al and bl included in Figure 2, and Figures al, bl, cl and dl included in Figure 4 show the images produced by the two video cameras depicted in Figure 1, on the one hand, and by the four video cameras depicted in Figure 3, on the other.

[0021] Using the signals corresponding to the abovementioned images, the unit 1 is able, in accordance with known principles (typically by using known image processing algorithms), to extract respective sets of data representing the abovementioned shapes of the bodies present within the region S. For example, Figures a2 and b2 of Figure 2, and Figures a2, b2, c2 and d2 of Figure 3 show the shapes, marked C of the body A present within the region S corresponding to images a1 and b1, or a1, b1, c1 and d1, produced, from their respective points of observation, by the video cameras 2.

[0022] For an overview of these algorithms, the following may usefully be referred to: Huang T.S., "Image sequence processing and dynamic scene analysis", Springer-Verlag, 1982; Jain A., "Fundamentals of digital image processing", Prentice Hall, 1989; Jain J.R. and Martin W.N. and Aggarwaal J.K., "Segmentation through the detection of changes due to motion", Computer Graphics and Image Processing, vol. 11, pp. 13-34, 1979; Debuisson M.-P., "Contour extraction of moving objects in complex outdoor scenes", Int. Journal of Computer Vision, vol. 14, pp.83-105, 1995; Bichsel M., "Segmenting Simply Connected Moving Objects in a Static Scene", IEEE Transactions on Pattern Analysis

and Machine Intelligence, vol. 16, n. 11, pp.1138-1142, 1994, as well as to US-A-5,212,547 or US-A-5,877,804.

[0023] In particular it is possible to have the abovementioned shapes correspond only to bodies possessed of movement in themselves, thus eliminating - for the purposes of the subsequent processing - information corresponding to fixed parts of the image, such as the outlines of the region S represented by broken lines in parts a2 and b2, and a2, b2, c2 and d2 in Figures 2 and 4, or to the object B. As an example, it may be imagined that the region S is a room in a museum, the body A is the body of a visitor moving about in the room, and the body B is a bench situated - in a fixed position of course - in the centre of the room.

[0024] In particular, it is known that objects having their own movement can be distinguished from fixed objects/items in such a way as to avoid, on the one hand, deception relating to the presence of an intruder body attempting to evade detection by moving very slowly and, on the other hand, the generation of false objects connected with, for example, vibratory movements, draughts, etc.

[0025] In one embodiment of the invention, the processing unit 1 can be programmed to find matches on the images between projections of the same real-world points belonging to the detected bodies. In this way, information is extracted from each image frame about the positions of the projections of the same real-world points onto the different image planes corresponding to the different observation points identified by the video cameras 2 (or equivalent image signal generating means). In particular it will be realized that, other parameters being equal, the availability of a larger number of observation points, and hence of a larger number of images (e.g. four rather than two) gives a certain degree of redundancy which can be used to increase the accuracy and reliability of the detection action.

[0026] Given a knowledge of the intrinsic and extrinsic parameters of the video cameras 2 - i.e. the focal distances of each lens fitted to the video cameras, the resolution of the sensor of each, the alignment of the optical axis with the centre of the active sensor of the video camera, the spatial co-ordinates relative to the origin of a reference system and the inclinations with respect to the axes of the reference system - the three-dimensional spatial positions of all observed points are unambiguously determined.

[0027] From these positions it is possible, with known methods, to extract the shape of the external surface of the objects present within the region S or an approximation thereto and from this to derive the volumetric map.

[0028] For an overview of the abovementioned methods the following may usefully be referred to: Boll R.M. and Vemuri B.C., "On three dimensional surface reconstruction methods", IEEE Transactions on Pattern Analysis and Computer Intelligence, vol. 13, no. 1, pp. 1-13, 1991; Besl P. and Jain R., "Three dimensional object recognition", Comp. Surveys, vol. 17, pp. 75-145, 1985; Aggarwaal J.K. et al, "Survey: representation methods of three dimensional objects", Progress in Pattern Recognition, vol. 1, North Holland, 1981; Morasso P. and Sandini G., "3D reconstruction from multiple stereo views", Proceedings 3rd International Conference on Image Analysis and Processing, 1985.

[0029] Using these techniques there may be uncertainty in the identification of surface parts of an object that are not seen simultaneously by at least two video cameras. This problem can be corrected by selecting a suitable location for another video camera or drawing on a priori information about the shape of the observed objects.

[0030] In all cases the characteristics of volume, shape, position and dimensions of the bodies in question can be represented by values associated with a finite set of parameters P.

[0031] Figure 5 shows the currently preferred method of producing the volumetric map of bodies in the monitored region S.

[0032] More specifically, the processing unit 1 is programmed, for the purposes of processing the signals generated by the video cameras 2, to divide up the entire volume of the region S into volumetric cells of fixed dimensions (some of these are shown diagrammatically at D1 and D2 in Figure 5), each cell corresponding to a portion of the real space inside the region S. For example, if the region S is a room in a museum, it may be decided to define the cells in question as cubic volumes with sides of, for example, ten centimetres. However, this is not of course a limiting option as it has to do with the spatial resolution with which it is wished to detect and locate the objects: the smaller the cells, the greater the resolution and vice versa.

[0033] Using a perspective projection function, each three-dimensional cell is projected onto the image plane of each video camera. In general, therefore, for each cell there is a certain area E on each image, as shown at the top of Figure 5, the bottom part of which meanwhile shows the location of the cells D1 and D2 within the three-dimensional Cartesian reference system used for locating the cells in question within the region S. Cells corresponding to regions of volume that are not covered by at least two video cameras are ignored.

[0034] The volumetric map F (see Figure 6) is obtained by checking each cell to see whether the areas corresponding to that cell's projection on the different images represent some portion of the objects present within the monitored region. If the outcome of the check is positive, that cell is judged to be occupied (for example, D1 is an example of this); otherwise it is judged to be unoccupied, as in the case of the cells marked D2. In this way the set of all occupied cells in each frame provides information about the volumetric occupation of the monitored region. From this information a volumetric map of occupation can be constructed almost immediately by checking which and how many of the cells into

which the monitored space has been divided are occupied by objects. The aim of all this is to obtain, as a result, the representation seen in Figure 6. Using this technique, the volumetric map approximates to the volume of actual occupation with a resolution based on the dimensions of the cells D1, D2 and on the spatial resolution of the means employed to generate the image signals (in particular, in the case of video cameras composed of a matrix of sensitive points, the resolution in a particular direction is given by the ratio of the dimension of the observed region to the number of sensitive elements in that direction). The dimensions of the cells are defined having regard not only to the resolution but also to the processing capacity of the unit 1 and of the frequency with which the surveillance information is to be updated. Also borne in mind is the fact that the overall degree of approximation can be enhanced, as already stated, by using more video cameras in suitable positions.

**[0035]** In particular, the processing unit 1 can be programmed, again in a known manner, to carry out a volumetric analysis of bodies for the purpose of recognizing the distinctive characteristics of shape and volume of the objects analysed. The programming can be done by conventional algorithmic approaches, thus coding ab initio the characteristics of shape and volume of the bodies to be detected into the processing system, or using statistical learning techniques such as for example neural networks. It is also possible to design the unit 1 such that it is able to evaluate the way the positions of the bodies examined within the region S are changing in space and time and deduce the dynamics of their movements, in particular the line, and the direction along this line, of their displacements.

**[0036]** This point will become clearer on referring to the flow diagram given in Figure 7, which shows, in a deliberately schematic way, for ease of comprehension, the principles by which the functions of automatic monitoring are carried out in the unit 1.

**[0037]** Assuming the process to start at a starting step 100, in a step 101 the unit 1 examines the data set corresponding to the images generated by the video cameras 2 (optionally already processed to refer only to moving objects) and in a second step 102 commences an action of scanning the region S in such a way as to scan the cells D into which the region S has been theoretically divided up. As a general rule, each of these cells will be identified by three co-ordinates $x_i$, $y_i$, $z_i$ within the system x, y and z to which the bottom part of Figure 5 refers.

**[0038]** From now on it will be assumed, for simplicity, that this scanning operation applies, on each successive detection of the images of the region S, to all the cells contained within the region S scanned on a "matrix" principle, for example in successive lines (co-ordinate x), columns (co-ordinate y) and planes (co-ordinate z).

**[0039]** Those skilled in the art of image processing will have realized that it is possible (e.g. in order to reduce the processing cost and/or speed up the processing) to adopt different scanning systems, such as predictive-type scanning systems which, once initialized with reference to a map of initial volumetric occupation, perform subsequent scans only on cells where there exists some degree of likelihood inherent in the fact that these cells may be significant in the generation of subsequent maps, the aim being to avoid the need to perform exhaustive scanning of the entire region S for each updating operation.

**[0040]** In this context it is also known that it is possible to intervene in such a way that, when operating on the abovementioned principles, the unit 1 is also capable of detecting, for example, the entry into the region S of a body not previously present, the aim being to extend the scanning action to those cells (previously not included in the scanning action) which the body subsequently occupies.

**[0041]** The steps marked 1031, 1041; 1032, 1042; ...; 103n, 104n indicate successive processing stages, here shown as carried out in parallel, though in fact they can be performed serially, and therefore sequentially in time. In the course of these steps, for each video camera 1, ..., n (n is equal to 2, and to 4, in the illustrative embodiments shown in Figures 1 and 3, respectively) and for each cell $D(x_i, y_i, z_i)$ that is scanned, the unit 1 checks to see whether the cells corresponding to their respective images generated by the video cameras 2 can be regarded as occupied or unoccupied.

**[0042]** In the next step, indicated by the general reference 105, the results of the comparisons carried out in steps 1041, 1042, ..., 104n are processed in order to decide whether, on the basis of the image data, the scanned cell is to be regarded as occupied or unoccupied for the purposes of constructing the map of volumetric occupation.

**[0043]** The relevant criteria for attributing the "occupied" or "unoccupied" logic value may differ.

**[0044]** On this subject it should be remembered that the cells of the region S are not necessarily all covered by all of the video cameras 2. As a consequence, in the case of certain cells, attribution of the "occupied" value may be based on a different number of decision processes relating to the individual images than the number of images taken into consideration in attributing the "occupied" logic value to other cells.

**[0045]** The criterion used in attributing the logic value in question may be of unanimous type (the cell is judged to be occupied for the purposes of the construction of the map of volumetric occupation if and only if all the video cameras 2 whose images are taken into account produce data corresponding to occupation in the relevant image), majority type (the cell is judged to be occupied if the majority of video cameras 2 give data indicating occupation in the respective images), or correlation with the values attributed to adjacent cells (so that uncertainty in the attribution of the "occupied" value to a cell is resolved on the basis of confident values attrib-

uted to spatially adjacent cells) or different again, according to well-known criteria in the image processing field.

**[0046]** The step 106 in Figure 7 represents simply the selection step where it is decided whether or not the scan of the region S (or of the scanned subregion thereof) can be said to be complete.

**[0047]** If the answer to this is negative, the process returns upstream of the step 102 and another cell is analysed.

**[0048]** If the result of the comparison in step 106 is positive, this indicates that the map of volumetric occupation is complete. At this point the map itself, which can be represented as illustrated diagrammatically in Figure 6 (which must of course be understood to be a perspective representation of a data set which in reality is three-dimensional), is subjected to a processing step 107 for the extraction of a set of parameters P which represent in compact form the shape, volume, position and/or dimensions of the detected bodies. By using algorithms that search for connected regions, it is possible to separate out elements of the volumetric map which are associated with different bodies; by this means it is possible to derive a volumetric map for every detected body. As a rule, for each volumetric map the characteristic parameters can be found by using automatic methods that are known in themselves. For instance, the map of volumetric occupation, that is the set of occupied cells and their positions in space, can be used directly as the volume parameter, the position of the centre of mass of each volumetric map can be used to represent the position of the body, and the dimensions of width, length and height of the smallest parallelepiped which inscribes the occupied volume can be used as dimensional parameters.

**[0049]** For an overview of some of the abovementioned methods, the following may usefully be referred to: Requicha A.G., "Representation of rigid solids: theory, methods and systems", Comp. Surveys, vol. 12, pp. 437-464, 1980; Requicha A.G. and Rossignac J.R., "Solid modeling and beyond", IEEE Computer Graphics and Applications, vol. 12, pp. 31-44, 1992; Aggarwaal J.K. and Cai Q., "Human motion analysis: a review", Proceedings of IEEE Computer Society Workshop on Motion of Non-Rigid and Articulated Objects, pp. 90-102, 1997.

**[0050]** In the next step 108 at least one of the parameters P obtained in this way is compared with a predetermined "model". The purpose of this is to establish whether or not the map F, corresponding to the position, size and shape of a body such as the body C, is "compatible" with the criteria of monitoring or surveillance which the system according to the invention has to follow.

**[0051]** One possible model for comparison may correspond to a defined part of the region of space S in which the body C must come no closer than a limiting distance. In this case compatibility is checked by using, for example, the parameters P relating to the position and dimensions of the detected bodies.

**[0052]** To take a concrete example, in Figure 6 the volume D corresponding to the region of space S which the body C must not enter may be an area that must be respected around a work of art exhibited in a museum (e.g. a picture hanging on a wall). To take another example, such as industrial equipment, the volume D may be a zone that must be respected around a machine with moving parts or with exposed parts at a high temperature and/or voltage.

**[0053]** In practice, in step 108 the unit checks (by applying known criteria) that, for example, none of the cells contained within the map of volumetric occupation F falls inside the volume D or is at a distance less than a minimum safety distance from the volume D.

**[0054]** If this condition is not found, so that the map F is compatible with the abovementioned model (to refer to the examples discussed above: the visitor has kept away from the picture hanging on the wall or the machine operator has kept a safe distance from the dangerous machine), the unit 1 prepares itself to repeat the monitoring action with reference to the next set of images taken by the video camera 2. The processing action thus returns upstream of step 101.

**[0055]** If, however, the map F is found to be incompatible with the model (for example because the visitor is found to have moved too close to the picture, or the machine operator has moved too close to the dangerous machine), the processing action moves on from step 108 to a new step 109 corresponding to the emission of a warning signal. This may be represented by e.g. an acoustic or visual alarm signal (optionally at a distance, aimed at a manned remote control station) emitted by a corresponding device 3. The device 3 must be understood to be of known type, depending on the alarm signal which it is wished to produce: it may for example be a siren, an acoustic indicator, a remote warning system, etc., connected to the unit 1.

**[0056]** From the above description it will be clear that by using the volumetric map describing the objects present in the monitored region S and obtained for example by the means described above or by the equivalent methods, and the manner in which it changes over time, it is possible to derive a description of the shapes of the objects and of their movements within the region S by encoding the information in numerical strings which describe at least one of the values of the characteristics of position, shape, dimensions and volume. The volumetric map of each body detected inside the monitored region and/or the manner in which it changes while the bodies are present in the monitored region can be compared with models of volumetric maps for other bodies, encoded in a similar manner and previously stored in the processing unit (take for example the model marked D in Figure 6) in order to recognize those bodies which must be detected from among all the bodies present inside the monitored region. The bodies may for example

be people only.

[0057] Furthermore, it is possible to detect the simultaneous presence of several bodies, even if of different kinds, in the monitored region. The manner in which the position of the bodies change within the monitored region can be used to detect violation of predefined subregions. It is thus possible to monitor, as has already been seen, the presence of a movement of people in the vicinity of a machine in an industrial environment and activate an alarm signalling procedure whenever at least one person comes within a certain distance of that machine.

[0058] The solution described is highly robust and overcomes the functional limitations of currently used systems. Thus, it is capable of detecting the presence and at the same time determining the position of people or objects within a defined region of space, discriminate between objects and people, between objects or people close to each other, and between objects and people that move into the monitored region following different paths or more generally with behaviours which could easily deceive other types of sensor.

[0059] Those skilled in the art will recognize that the method according to the invention can be carried out using, at least in part, a computer program capable of being run on a computer in such a way that the system comprising the program and the computer carries out the method according to the invention. The invention therefore extends also to such a program capable of being loaded into a computer which has the means of or is capable of carrying out the method according to the invention, as well as to the corresponding information technology product comprising a means readable by a computer containing codes for a computer program which, when the program is loaded into the computer, cause the computer to carry out the method according to the invention.

[0060] Clearly, without affecting the principle of the invention, the constructional details and the embodiments may be greatly altered compared to what has been described and illustrated, without thereby departing from the scope of the present invention, as defined in the accompanying claims.

## Claims

1. Method for the detection and location of bodies (A, B) in a defined region of space (S), comprising the operations of generating (2) image signals capable of representing a succession of images of at least one body present in the said region (S), each image corresponding to a defined instant, characterized in that it comprises the following operations:

   - processing (101 to 106) the said image signals in such a way as to obtain for each instant taken into consideration a volumetric map (F) of the said at least one body present in the said region (S), the said volumetric map (F) representing the shape, position, volume and dimensions of the body to which the said volumetric map (F) refers,
   - extracting (107) from the said volumetric map (F) at least one parameter (P) taken from the following group: descriptors of shape and volume, such as the volumetric map (F) itself, the position co-ordinates and the dimensions of the said at least one body to which the said volumetric map (F) refers,
   - comparing (108) the said at least one parameter (P) with at least one model (D) for compatibility of the said volumetric map (F) with predetermined conditions of occupation of the said region (S), and
   - selectively generating a warning signal (109) depending on the outcome of the said comparison (108).

2. Method according to Claim 1, characterized in that it comprises the operations of storing volumetric maps (F) or successions of the said at least one parameter (P) obtained from image signals relating to images of the said succession corresponding to successive instants, in order to detect changes in time in the said volumetric map (F) or the said at least one parameter (P), and in that the said model is itself generated as a model of changes over time.

3. Method according to Claim 1 or Claim 2, characterized in that it comprises the operation of comparing (108) successions of the at least one parameter (P) obtained from image signals relating to images of the said succession corresponding to successive instants with at least one model for compatibility of said successions with predetermined conditions of occupation of the said region (S).

4. Method according to Claim 1 or Claim 2, characterized in that it comprises the following operations:

   - generating image signals relating to the view of the said region (S) from separate observation points (2) so as to generate at least two separate image signals relating respectively to the projections of the same points of the said region (S) viewed from separate observation points,
   - processing (1) the said separate image signals by finding the match between the projections of the same real-world points onto separate images with a view to finding its position in space,
   - obtaining the said volumetric map (F) from the said positions in space.

5. Method according to Claim 1 or Claim 2, characterized in that it comprises the following operations:

- subjecting at least part of the said region (S) to scanning of cells ($D(x_i, y_i, z_i)$),
- detecting (1041, 1042, ..., 104n) for each of the said cells the said separate signals (1031, 1032, ..., 103n),
- generating (105) for each scanned cell and from the values of the said separate signals detected for the said cell, an occupation signal whose value can identify the occupation of the scanned cell by the said at least one body present in the said region of space (S), the said volumetric map (F) being identified by the set of values attributed to the said occupation signal corresponding to the scanned cells.

6. Apparatus for the detection and location of bodies (A, B) in a defined region of space (S) comprising image signal generating means (2) capable of monitoring the said region (S) in order to produce an image or succession of images of at least one body present in the said region (S), each image corresponding to a defined instant, the apparatus being characterized in that it comprises processing means (1) configured (101 to 106) so as to:

   - process the said image signals in such a way as to obtain for each instant taken into consideration a volumetric map (F) of the said at least one body present in the said region (S), the said volumetric map (F) representing the shape, position, volume and dimensions of the body to which the said volumetric map (F) refers,
   - extract (107) from the said volumetric map (F) at least one parameter (P) taken from the following group: descriptors of shape and volume, such as the volumetric map (F) itself, the position co-ordinates and the dimensions of the said at least one body to which the said volumetric map (F) refers,
   - compare (108) the said at least one parameter (P) with at least one model (D) for compatibility of the said volumetric map (F) with predetermined conditions of occupation of the said region (S),
   - and in that it also comprises warning means (3) connected to the said processing means (1) for selectively generating a warning signal (109) depending on the outcome of the said comparison (108).

7. Apparatus according to Claim 6, characterized in that the said processing means (1) are configured so as to store volumetric maps (F) or successions of the said at least one parameter (P) obtained from image signals relating to images of the said succession corresponding to successive instants, in order to detect changes in time in the said volumetric map (F) or the said at least one parameter (P), and in that the said model is itself generated as a model of changes over time.

8. Apparatus according to Claim 6 or Claim 7, characterized in that the said processing means (1) are configured in such a way as to compare (108) successions of the at least one parameter (P) obtained from image signals relating to images of the said succession corresponding to successive instants with at least one model for compatibility of said successions with predetermined conditions of occupation of the said region (S).

9. Apparatus according to Claim 6 or Claim 7, characterized in that:

   - a plurality of means (2) are provided for the generation of image signals relating to views of the said region (S) from separate observation points (2) so as to generate at least two separate image signals relating respectively to the projections of the same points of the said region (S) viewed from separate observation points,
   - the said processing means (1) are configured so as to process the said separate image signals by finding the match between the projections of the same real-world points onto separate images in order to find its position in space and obtain the said volumetric map (F) from the said positions in space.

10. Apparatus according to Claim 6 or Claim 7, characterized in that the said processing means (1) are configured so as to:

   - subject at least part of the said region (S) to scanning of cells ($D(x_i, y_i, z_i)$),
   - detect (1041, 1042, ..., 104n) for each of the said cells the said separate signals (1031, 1032, ..., 103n),
   - generate (105) for each scanned cell and from the values of the said separate signals detected for the said cell, an occupation signal whose value can identify the occupation of the scanned cell by the said at least one body present in the said region of space (S), the said volumetric map (F) being identified by the set of values attributed to the said occupation signal corresponding to the scanned cells.

11. Computer program capable of being run on a computer in such a way that the system comprising the program and the computer carries out the method according to any one of Claims 1 to 5.

12. Computer program capable of being loaded into a computer which has the means of or is capable of carrying out the method according to any one of

Claims 1 to 5.

13. Information technology product comprising a means readable by a computer containing codes for a computer program which, when the program is loaded into the computer, cause the computer to carry out the method according to any one of Claims 1 to 5.

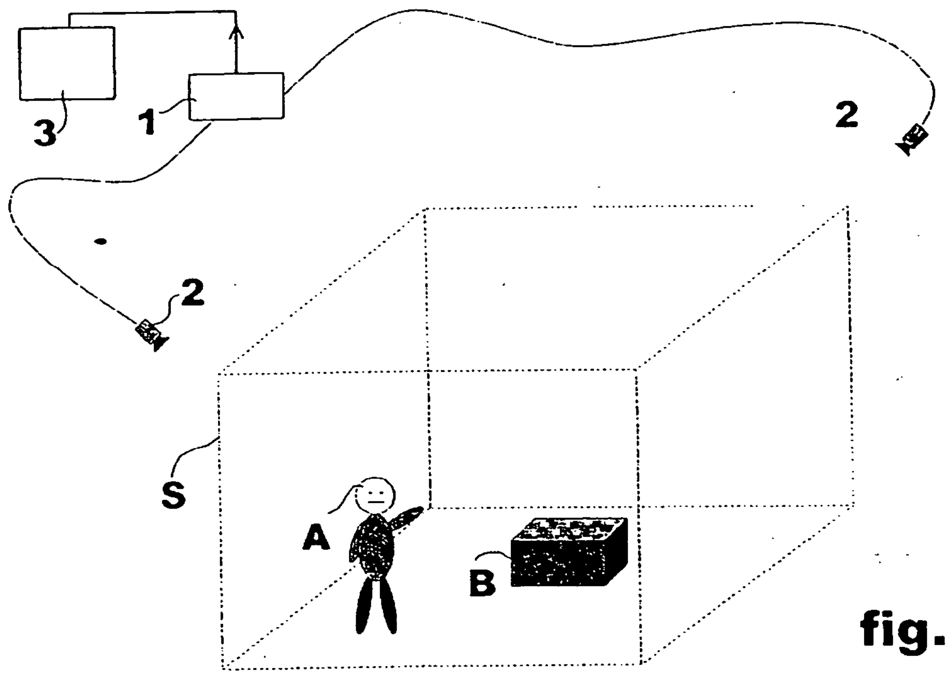
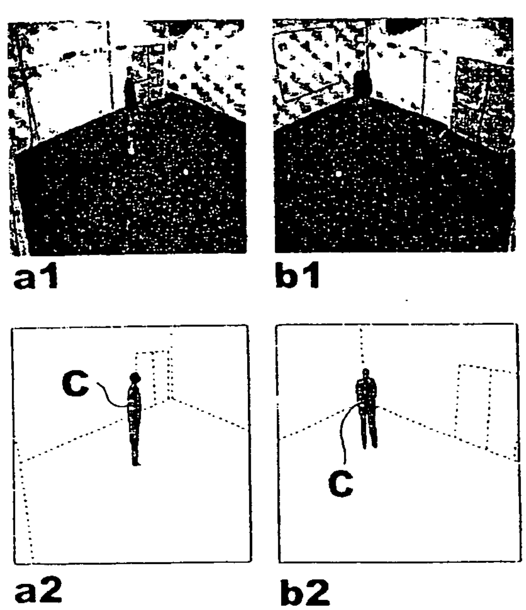



fig.1

fig.2



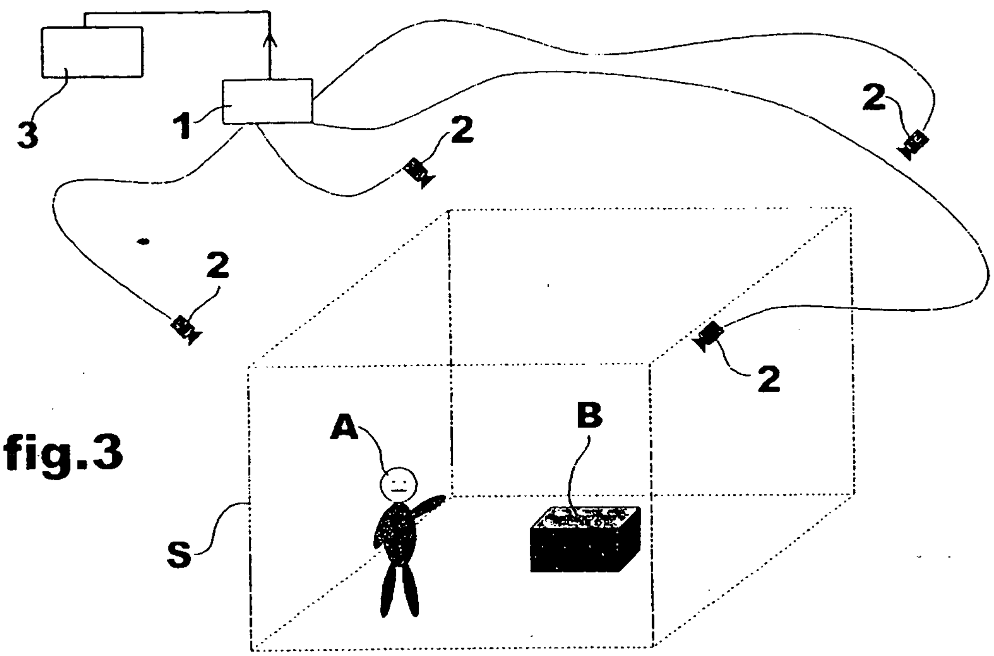
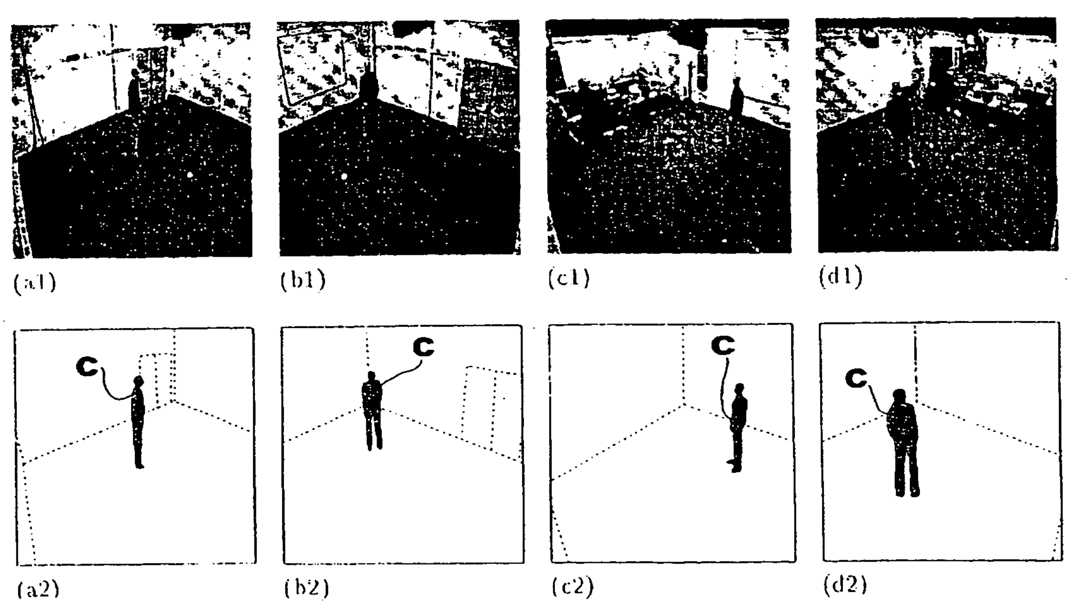



fig.3

fig.4



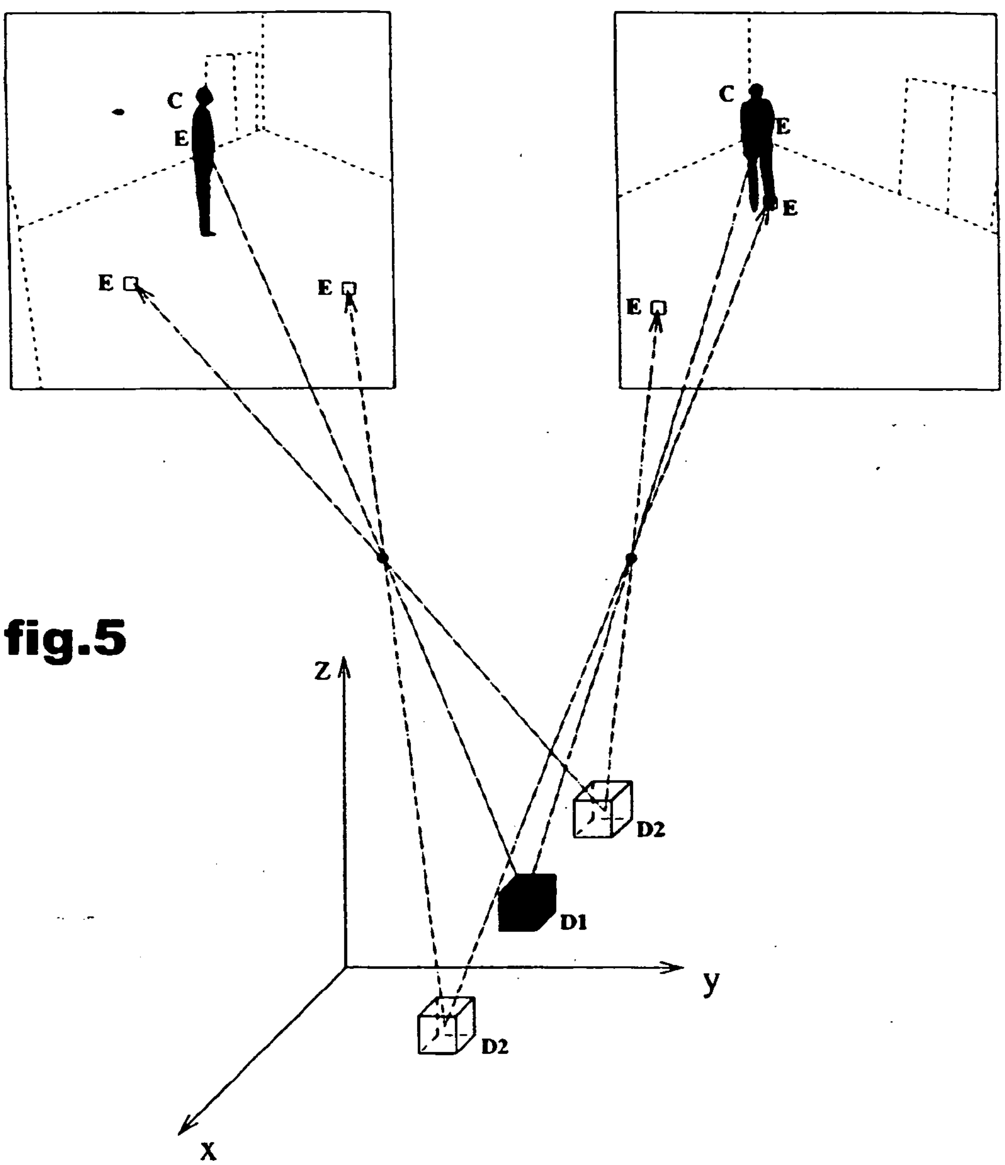



fig.5

fig.6

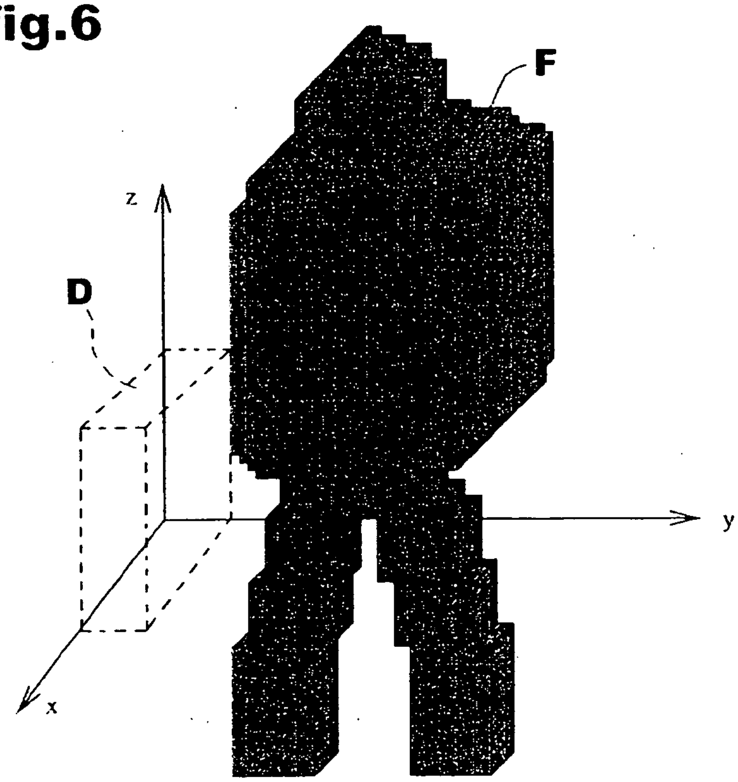
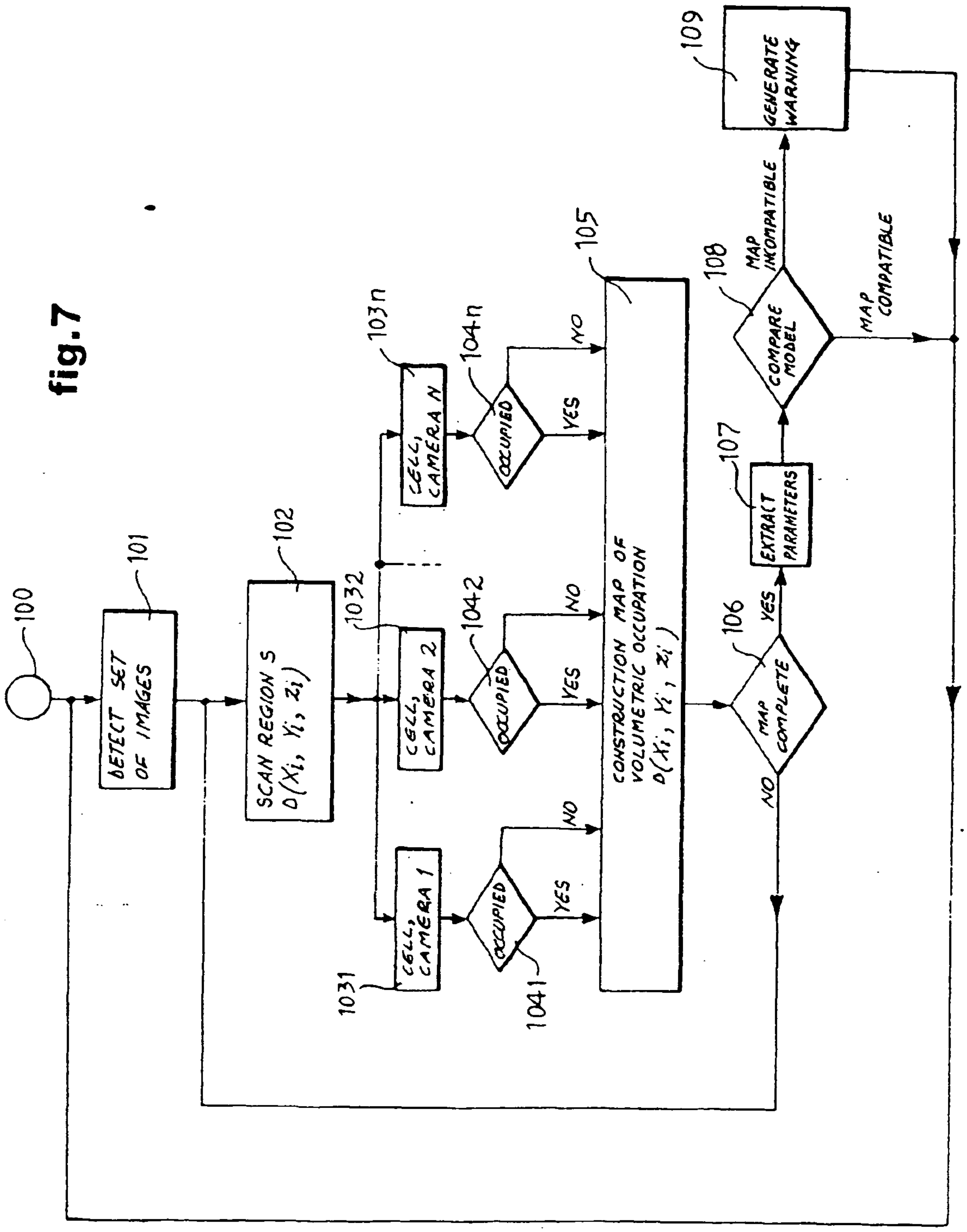


fig.7


100
101
DETECT SET OF IMAGES
102
SCAN REGION S D(Xi, Yi, Zi)
1031
CELL, CAMERA 1
1032
CELL, CAMERA 2
103n
CELL, CAMERA N
1041
1042
104n
OCCUPIED
YES
NO
105
CONSTRUCTION MAP OF VOLUMETRIC OCCUPATION D(Xi, Yi, Zi)
106
MAP COMPLETE
107
EXTRACT PARAMETERS
108
COMPARE MODEL
MAP INCOMPATIBLE
MAP COMPATIBLE
109
GENERATE WARNING

European Patent Office

# EUROPEAN SEARCH REPORT

Application Number
EP 99 83 0376

## DOCUMENTS CONSIDERED TO BE RELEVANT

| Category | Citation of document with indication, where appropriate, of relevant passages | Relevant to claim | CLASSIFICATION OF THE APPLICATION (Int.Cl.7) |
|---|---|---|---|
| Y | EP 0 356 734 A (SIEMENS AKTIENGESELLSCHAFT) 7 March 1990 (1990-03-07) * column 2, line 53 - column 3, line 5 * | 1,6 | G08B13/194 |
| Y | TARBOX G H ET AL: "VOLUMETRIC BASED INSPECTION*" PROCEEDINGS OF THE IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS,US,NEW YORK, IEEE, vol. -, page 1239-1246 XP000334081 ISBN: 0-7803-0738-0 * abstract * | 1,6 | |
| A | CARLSON J ET AL: "Real-time 3D visualization of volumetric video motion sensor data" SURVEILLANCE AND ASSESSMENT TECHNOLOGIES FOR LAW ENFORCEMENT, BOSTON, MA, USA, 19-20 NOV. 1996, vol. 2935, pages 69-79, XP000863382 Proceedings of the SPIE - The International Society for Optical Engineering, 1997, SPIE-Int. Soc. Opt. Eng, USA ISSN: 0277-786X * abstract * | 1,6 | |

TECHNICAL FIELDS SEARCHED (Int.Cl.7)
G08B

The present search report has been drawn up for all claims

| Place of search | Date of completion of the search | Examiner |
|---|---|---|
| THE HAGUE | 16 December 1999 | Chateau, J-P |

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

## ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 99 83 0376

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

16-12-1999

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---|---|---|
| EP 356734 A | 07-03-1990 | NONE | |

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82